

**PEMODELAN REGRESI LINIER MULTIVARIAT DENGAN METODE  
PEMILIHAN MODEL *FORWARD SELECTION* DAN *ALL POSSIBLE  
SUBSET SELECTION* PADA JUMLAH KEMATIAN BAYI DAN INDEKS  
PEMBANGUNAN MANUSIA (IPM)  
( Studi Kasus di Provinsi Jawa Tengah Tahun 2013 )**



---

**SKRIPSI**

---

**Disusun Oleh :  
INDRI PUSPITASARI  
NIM 24010211140088**

**JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG  
2015**

**PEMODELAN REGRESI LINIER MULTIVARIAT DENGAN METODE  
PEMILIHAN MODEL *FORWARD SELECTION* DAN *ALL POSSIBLE  
SUBSET SELECTION* PADA JUMLAH KEMATIAN BAYI DAN INDEKS  
PEMBANGUNAN MANUSIA (IPM)  
( Studi Kasus di Provinsi Jawa Tengah Tahun 2013 )**

**Oleh :  
INDRI PUSPITASARI  
NIM 24010211140088**

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar  
Sarjana Sains pada Jurusan statistika**

**JURUSAN STATISTIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG  
2015**

## HALAMAN PENGESAHAN I

**Judul : Pemodelan Regresi Linier Multivariat dengan Metode Pemilihan Model *Forward Selection* dan *All Possible Subset Selection* pada Jumlah Kematian Bayi dan Indeks Pembangunan Manusia (IPM) (Studi Kasus di Provinsi Jawa Tengah Tahun 2013)**

**Nama : Indri Puspitasari**

**NIM : 24010211140088**

Telah diujikan pada sidang Tugas Akhir tanggal 28 Agustus 2015 dan dinyatakan lulus pada 25 September 2015.

Semarang, 28 September 2015

Mengetahui,

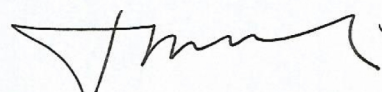
Sekretaris Jurusan Statistika

Fakultas Sains dan Matematika UNDIP

Panitia Penguji Ujian Tugas Akhir

Ketua,

  
**Drs. Agus Rusgivorfo, M.Si**  
**NIP. 196408131990011001**

  
**Dra. Dwi Ispriyanti, M.Si**  
**NIP. 195709141986032001**

## HALAMAN PENGESAHAN II

**Judul : Pemodelan Regresi Linier Multivariat dengan Metode Pemilihan Model *Forward Selection* dan *All Possible Subset Selection* pada Jumlah Kematian Bayi dan Indeks Pembangunan Manusia (IPM) (Studi Kasus di Provinsi Jawa Tengah Tahun 2013)**

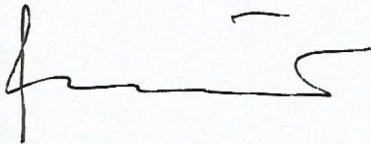
**Nama : Indri Puspitasari**

**NIM : 24010211140088**

Telah diujikan pada sidang Tugas Akhir tanggal 28 Agustus 2015 dan dinyatakan lulus pada 25 September 2015.

Semarang, 28 September 2015

Pembimbing I



Abdul Hoyyi, S. Si, M.Si

NIP. 197202022008011018

Pembimbing II



Diah Safitri, S.Si, M.Si

NIP. 197510082003122001

## KATA PENGANTAR

Puji Syukur penulis ucapkan atas kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penulisan Tugas Akhir ini dengan judul “**Pemodelan Regresi Linier Multivariat dengan Metode Pemilihan Model *Forward Selection* dan *All Possible Subset Selection* pada Jumlah Kematian Bayi dan Indeks Pembangunan Manusia (IPM) (Studi Kasus di Provinsi Jawa Tengah Tahun 2013 )**”.

Penulis menyadari bahwa dalam penulisan laporan ini tidak lepas dari bimbingan dan dukungan yang diberikan beberapa pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada :

1. Ibu Dra. Dwi Ispriyanti, M.Si sebagai Ketua Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro.
2. Bapak Abdul Hoyyi, S.Si.,M.Si dan Ibu Diah Safitri, S.Si, M.Si selaku dosen pembimbing I dan dosen pembimbing II
3. Bapak dan Ibu dosen Jurusan Statistika Fakultas Sains dan Matematika Universitas Diponegoro
4. Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu penulis dalam penulisan Laporan Tugas Akhir ini.

Penulis menyadari bahwa dalam penulisan tugas akhir ini masih jauh dari sempurna. Oleh karena itu kritik dan saran yang membangun sangat penulis harapkan demi perbaikan dalam kesempatan berikutnya.

Semarang, September 2015

Penulis

## ABSTRAK

Analisis regresi merupakan analisis statistik yang bertujuan untuk mengukur pengaruh variabel dependen terhadap variabel independen. Regresi Linier Multivariat adalah regresi dengan model yang terdiri dari lebih dari satu variabel dependen dan antar variabel dependen saling berkorelasi. Jumlah Kematian Bayi dan Indeks Pembangunan Manusia (IPM) di Provinsi Jawa Tengah tahun 2013 dipengaruhi oleh beberapa variabel, antara lain rata-rata lama sekolah dan jumlah puskesmas. Untuk menganalisis pengaruh tersebut digunakan analisis regresi linier multivariat karena antar variabel dependen saling berkorelasi. Pemilihan model ditentukan dengan menggunakan *Forward Selection* dan *All Possible Subset Selection*. Pemilihan model dengan menggunakan *Forward Selection*, variabel yang pertama dimasukkan dalam model didasarkan pada variabel independen yang memiliki korelasi terbesar dengan variabel dependen. Pada metode *All Possible Subset Selection*, pemilihan model dilakukan dengan memodelkan semua model yang mungkin terbentuk. Kriteria AIC digunakan dalam penentuan model pada metode *All Possible Subset Selection*. Diperoleh model menggunakan metode *Forward Selection* dan *All Possible Subset Selection* adalah model yang memiliki variabel independen yang sama, yaitu model dengan variabel independen rata-rata lama sekolah dan jumlah puskesmas. Error dari model tersebut memenuhi semua asumsi error. Berdasarkan model tersebut, diperoleh nilai AIC sebesar 247,8142 dan *Eta Squared Lambda* sebesar 92,22 %.

**Kata Kunci** : Regresi Linier Multivariat, Forward Selection, All Possible Subset Selection, AIC, IPM, Jumlah Kematian Bayi, Jawa Tengah

## ***ABSTRACT***

Regression analysis is a statistical analysis that aims to measure the effect of the independent variables to the dependent variable. Multivariate Linear Regression is a regression model that consists of more than one dependent variables and the dependent variables are correlated. The Number of Infant Mortality and Human Development Index (HDI) of Central Java Province in 2013 was influenced by several variables, such as: mean years of schooling and the number of health centers. To analyze the effects of mean years of schooling and the number of health centers to The Number of Infant Mortality and Human Development Index (HDI) can use multivariate linear regression analysis because the dependent variables are correlated. Model selection is determined by using the Forward Selection and All Possible Subset Selection. Selection the model by using Forward Selection, first variables that is included in the model is based of independent variable that have the greatest correlation with the dependent variables. For All Possible Subset Selection, model selection is done by modeling all the models that may have formed. AIC criteria is used for determining the model for All Possible Subset Selection. The model which is selected by using Forward Selection and All Possible Subset Selection has the same independent variables, the model with independent variables mean years of schooling and the number of health centers. The error of the model fulfill all the assumptions of error. Based on the model, the value of AIC is 247.8142 and Eta Squared Lambda is 92.22%.

**Keywords** : Multivariate Linear Regression, Forward Selection, All Possible Subset Selection, AIC, IPM, Number of Infant Mortality, Central Java.

## DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
KATA PENGANTAR .....	iv
ABSTRAK .....	v
<i>ABSTRACT</i> .....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR .....	x
DAFTAR TABEL.....	xi
DAFTAR LAMPIRAN .....	xii
BAB I PENDAHULUAN	
1.1. Latar Belakang.....	1
1.2. Batasan Masalah.....	3
1.3. Perumusan Masalah.....	3
1.4. Tujuan Penelitian .....	4
BAB II TINJAUAN PUSTAKA	
2.1. Jumlah Kematian Bayi dan Indeks Pembangunan Manusia (IPM).....	5
2.1.1. Jumlah Kematian Bayi.....	5
2.1.2. Indeks Pembangunan Manusia (IPM).....	6
2.2. Distribusi Normal Multivariat .....	9
2.3. Korelasi .....	11
2.4. Pengujian <i>Independency</i> Antar Variabel .....	12
2.5. Analisis Regresi Linier .....	13
2.6. Analisis Regresi Linier Multivariat .....	13
2.7. Uji Signifikansi Model.....	19



2.8. Pengujian Subset X.....	20
2.9. Pemilihan Model.....	21
2.9.1. <i>Forward Selection</i> .....	21
2.9.2. <i>All Possible Subset</i> .....	22
2.10. Kriteria AIC .....	22
2.11. Hubungan Variabel Dependen dan Independen .....	23
2.12. Uji Asumsi Error.....	23
2.12.1. Normalitas.....	23
2.12.2. Homoskedastisitas.....	25
2.12.3. Nonautokorelasi .....	26
 BAB III METODOLOGI PENELITIAN	
3.1. Sumber Data.....	29
3.2. Variabel Penelitian .....	29
3.3. Langkah Metode Analisis Data.....	32
3.4. Diagram Alir Analisis ( <i>Flowchart</i> ).....	33
 BAB IV HASIL DAN PEMBAHASAN	
4.1. Deskripsi Data.....	36
4.2. Pengujian Variabel Dependen.....	37
4.2.1. Uji Normalitas Variabel Dependen.....	37
4.2.2. Pengujian Independency Antar Variabel .....	39
4.3. Model Awal Regresi Linier Multivariat.....	40
4.4. Model <i>Forward Selection</i> .....	41
4.5. Model <i>All Possible Subset Selection</i> .....	43
4.6. Uji Signifikansi Model.....	44
4.7. Uji Subset X.....	44
4.8. Uji Asumsi Error .....	45
4.8.1. Normalitas .....	45

4.8.2. Homoskedastisitas.....	47
4.8.3. Nonautokorelasi .....	48
4.9. Model Akhir .....	48
BAB V KESIMPULAN.....	50
DAFTAR PUSTAKA .....	51
LAMPIRAN.....	53

## DAFTAR GAMBAR

	Halaman
<b>Gambar 1.</b> Diagram Alir Analisis Data Penelitian .....	35
<b>Gambar 2.</b> Chisquared Plot Variabel Dependensi.....	37
<b>Gambar 3.</b> ChiSquared Plot Error.....	46
<b>Gambar 4.</b> DD-Plot Error .....	47

## DAFTAR TABEL

	Halaman
<b>Tabel 1.</b> Aturan Keputusan Uji Durbin-Watson.....	27
<b>Tabel 2.</b> Pemilihan Variabel Independen .....	30
<b>Tabel 3.</b> Korelasi Antar Variabel.....	31
<b>Tabel 4.</b> Variabel Penelitian .....	32
<b>Tabel 5.</b> Statistik Deskriptif Data Penelitian .....	36
<b>Tabel 6.</b> <i>Wilks Lambda</i> .....	42
<b>Tabel 7.</b> <i>Wilk's Lambda Partial</i> .....	42
<b>Tabel 8.</b> Nilai AIC .....	43
<b>Tabel 9.</b> Pengujian Subset X Model.....	45
<b>Tabel 10.</b> Uji Nonautokorelasi .....	48

## DAFTAR LAMPIRAN

	Halaman
<b>Lampiran 1.</b> Data Variabel Penelitian .....	53
<b>Lampiran 2.</b> Uji Normalitas Variabel Dependen .....	54
<b>Lampiran 3.</b> Uji <i>Bartlett of Sphericity</i> Variabel Dependen.....	56
<b>Lampiran 4.</b> Model Awal Regresi Linier Multivariat .....	57
<b>Lampiran 5.</b> Pemilihan Model Regresi Linier Multivariat dengan <i>Forward</i> <i>Selection</i> .....	58
<b>Lampiran 6.</b> Pemilihan Model Regresi Linier Multivariat dengan <i>All Possible Selection</i> .....	59
<b>Lampiran 7.</b> Model Regresi Linier Multivariat .....	61
<b>Lampiran 8.</b> Tabel Kolmogorov Smirnov .....	66
<b>Lampiran 9.</b> Tabel Chi Square .....	67
<b>Lampiran 10.</b> Tabel Wilks Lambda.....	68
<b>Lampiran 11.</b> Tabel Durbin Watson.....	69

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Dewasa ini, penurunan jumlah kematian bayi menjadi fokus pemerintah. Hal ini dikarenakan kedepannya anak akan berperan dalam proses pembangunan. Menurut Bappenas (2009), kelangsungan hidup anak sangat menentukan kualitas sumber daya manusia dimasa yang akan datang. Selain itu, Budiriyanto (2011) menyatakan bahwa kuantitas dan kualitas sumber daya manusia dibutuhkan dalam proses pembangunan. Dalam Unicef (2013), pemerintah berupaya untuk terus menurunkan jumlah kematian pada bayi dengan berbagai macam cara.

Menurut BPS (2011), baik buruknya pembangunan dapat dilihat dari nilai Indeks Pembangunan Manusia (IPM). IPM ini menentukan tingkat kinerja suatu daerah. Menurut BPS (2012), nilai IPM digunakan untuk melihat status pembangunan manusia dengan kriteria: Status pembangunan rendah ( $IPM < 50$ ), Status pembangunan menengah bawah ( $50 \leq IPM < 66$ ), Status pembangunan menengah atas ( $66 \leq IPM < 80$ ), dan Status pembangunan tinggi ( $IPM \geq 80$ ).

Berdasarkan Buku Profil Kesehatan Provinsi Jawa Tengah Tahun 2012, pada tahun 2012 jumlah kematian bayi adalah 6235 jiwa (Dinkes, 2013) dan menurut Buku Profil Kesehatan Provinsi Jawa Tengah Tahun 2013, pada tahun 2013 jumlah kematian bayi adalah 5865 jiwa (Dinkes, 2014). Diperoleh bahwa jumlah kematian bayi di Provinsi Jawa Tengah mengalami penurunan, yaitu sebesar 5,93 %. Menurut BPS (2014), pada tahun 2012 diperoleh IPM untuk Provinsi Jawa Tengah berada pada nilai 73,36, sedangkan pada tahun 2013 nilai

IPM untuk Provinsi Jawa Tengah berada pada nilai 74,05. Berdasarkan hasil tersebut diperoleh bahwa dari kurun waktu antara tahun 2012 sampai 2013 terjadi peningkatan pada nilai IPM untuk Provinsi Jawa Tengah yaitu sebesar 0,94 %.

Meskipun jumlah kematian bayi mengalami penurunan dan nilai IPM mengalami peningkatan, pemerintah harus berupaya untuk menuntaskan masalah tersebut. Maka diperlukan penelitian untuk mengetahui faktor-faktor yang dapat mempengaruhi penurunan jumlah kematian bayi dan peningkatan IPM. Berdasarkan paparan Bappenas (2009) dan BPS (2011), maka didapatkan variabel yang secara signifikan memiliki hubungan linier dengan jumlah kematian bayi dan IPM, yaitu rata-rata lama sekolah dan jumlah puskesmas. Pada variabel jumlah kematian bayi dan IPM, kedua variabel tersebut memiliki hubungan yang linier. Pada kasus seperti ini, metode analisis yang dapat dipakai untuk memodelkan jumlah kematian bayi dan IPM adalah analisis regresi linier multivariat. Menurut Rencher (2002), pada analisis regresi linier multivariat terdapat variabel dependen lebih dari satu variabel dan antar variabel dependen saling berhubungan. Jika terdapat variabel independen lebih dari satu maka penentuan model dapat menggunakan metode *Stepwise Procedure* dan metode *All Possible Subset Selection*. *Forward Selection* merupakan salah satu metode penentuan model dari *Stepwise Procedure* dimana dalam pemilihan model tersebut, variabel yang pertama dimasukkan dalam model didasarkan pada variabel independen yang memiliki hubungan linier terbesar dengan variabel dependen. Pada metode *All Possible Subset Selection* pemilihan model dilakukan dengan memodelkan semua model yang mungkin terbentuk. Untuk kriteria untuk pemilihan model dapat dilakukan menggunakan *Akaike's Information Criterion*

(*AIC*) (Johnson dan Winchern , 2007). Pada penelitian ini akan digunakan metode pemilihan model dengan menggunakan *Forward Selection* dan *All Possible Subset Selection*. Pada metode *All Possible Subset Selection* kriteria pemilihan model dengan menggunakan *AIC*.

Berdasarkan paparan tersebut, maka penulis melakukan penelitian untuk Tugas Akhir dengan judul “Pemodelan Regresi Linier Multivariat dengan Metode Pemilihan Model *Forward Selection* dan *All Possible Subset Selection* pada Jumlah Kematian Bayi dan Indeks Pembangunan Manusia (IPM) (Studi Kasus di Provinsi Jawa Tengah Tahun 2013)”.

## **1.2. Batasan Masalah**

1. Variabel bebas dalam penelitian ini adalah rata-rata lama sekolah dan jumlah puskesmas.
2. Pada metode *All Possible Subset Selection*, kriteria pemilihan model menggunakan *AIC*.

## **1.3. Perumusan Masalah**

Perumusan masalah dalam penelitian ini adalah bagaimana memodelkan jumlah kematian bayi dan IPM di Provinsi Jawa Tengah menggunakan analisis regresi linier multivariat dengan pemilihan model menggunakan *Forward Selection* dan *All Possible Subset Selection*.



#### **1.4. Tujuan Penelitian**

Tujuan penulis pada penelitian ini adalah memodelkan Jumlah Kematian Bayi dan IPM di Provinsi Jawa Tengah menggunakan analisis regresi linier multivariat dengan pemilihan model menggunakan *Forward Selection* dan *All Possible Subset Selection*.